R E P O R T R E S U M E S

ED 011 936 24 AA 000 053

AN ANALYTICAL APPROACH TO RESEARCH ON INSTRUCTIONAL METHODS.

BY- GAGE, N.L.

REPORT NUMBER BR-5-0252-2 PUB DATE 66
CONTRACT OEC-6-10-078
EDRS PRICE MF-$0.09 HC-$0.72 18P.

DESCRIPTORS- *TEACHER EDUCATION, *TEACHING, *RESEARCH, *BEHAVIOR, COMPREHENSION, *MICROTEACHING, STANFORD

THE APPROACH USED AT STANFORD UNIVERSITY TO RESEARCH ON TEACHING WAS DISCUSSED, AND THE AUTHOR EXPLAINED THE CONCEPTS OF "TECHNICAL SKILLS," "MICROTEACHING," AND "MICROCRITERIA" THAT WERE THE BASIS OF THE DEVELOPMENT OF THIS APPROACH TO RESEARCH AND TO STANFORD'S SECONDARY-TEACHER EDUCATION PROGRAM. THE AUTHOR PRESENTED A BASIC DISTINCTION BETWEEN RESEARCH ON LEARNING, WHICH HE TERMED AS "THE CONDITIONS UNDER WHICH A CHANGE OF BEHAVIOR DUE TO EXPERIENCE TAKES PLACE," AND RESEARCH ON TEACHING WHICH DEALS WITH THE CONDITIONS ESTABLISHED BY THE TEACHER UNDER WHICH LEARNING OCCURS IN THE PUPIL. THE FAILURE OF PAST RESEARCH ON TEACHING TO YIELD SIGNIFICANT RESULTS WAS CITED AS EVIDENCE OF A NEED FOR SOME MORE PRODUCTIVE APPROACH. THE NEW APPROACH DEVELOPED AT STANFORD REPLACES THE CONCEPT OF A SINGLE CRITERION OF TEACHER-EFFECTIVENESS BY AN ANALYSIS OF THE TEACHING PROCESS INTO TECHNICAL SKILLS, THE SPECIFIC INSTRUCTIONAL TECHNIQUES AND PROCEDURES THAT A TEACHER MAY USE IN THE CLASSROOM. BY THIS APPROACH, THE AUTHOR SAID MORE SATISFACTORY RESEARCH CAN BE DONE ON TEACHER EDUCATION AND TEACHER EFFECTS BY MEASURING INDEPENDENT AND DEPENDENT VARIABLES AND CONDUCTING TRUE EXPERIMENTS. THESE TECHNICAL SKILLS IN 5- TO 10-MINUTE PRACTICE-TEACHING SESSIONS. THE MICROCRITERION OF LECTURE EFFECTIVENESS, OR TEACHER TECHNICAL SKILLS IN 5-TO-10 MINUTE PRACTICE-TEACHING SESSIONS. THE MICRO-CRITERION OF LECTURE EFFECTIVENESS, OR TEACHER EXPLAINING-ABILITY, WAS MEASURED BY TESTS OF PUPIL COMPREHENSION FOLLOWING THE INTERN PRESENTATIONS. THE AUTHOR VIEWED THE TECHNICAL-SKILLS APPROACH AS A MEANS OF PROVIDING A SCIENTIFICALLY BASED SET OF ANSWERS TO THE QUESTION OF WHAT SHOULD A TEACHER DO IN THE CLASSROOM. (AL)

AA000053

BR-5-0252-2
OEC-6-10-078
P.A. 24

An Analytical Approach to Research on Instructional Methods

N. L. Gage
Stanford University

Instructional method constitutes one of the most important and promising but also the most frustrating of the areas of educational research and development. Compared with the areas of learning, subject matter, instructional materials, and organization for instruction, instructional method appeals to the classroom teacher as closest to the heart of her problem. It is all right for a teacher to know about learning, to know his subject matter, to have appropriate instructional materials, and to fit into a given organization for instruction. But what a teacher really wants to know is, "What should I do in the classroom?" If you ask prospective teachers or teachers on the job, "Where do you really want help?" I think the reply will deal with some aspect of instructional methods.

Unfortunately, the urgency of this demand has not been accompanied by corresponding success in meeting the demand. Research on teaching has yielded relatively few solid and usable results. The stature of theory and research in other areas puts them well ahead of the study of teaching in the struggle for scientific maturity.

Implicit in what I am saying is a basic distinction between research on learning and research on teaching. The former deals with all the conditions under which learning, or a change in behavior due to experience, takes place. And as I have already indicated, the study of learning is relatively mature, well established, with many volumes of substantial literature to its credit. Research on teaching, on the other hand, deals with a subset of the conditions under which learning occurs in one person, namely, the conditions established by the behaviors of another person, called the teacher. As our schools have

ED011936

developed during the past two or three millennia, we have always attempted to promote and improve the learning process through the intermediation of such another person. Until very recently, the assumption that teachers were helpful or even necessary for many important kinds of learning that society wanted to promote went unchallenged. Even today, the challenge of independent study or computer-assisted instruction and other devices is a mere whisper against the thunder of the assumption that teachers are necessary, that teachers are here to stay. My discussion of instructional methods is going to be based on that assumption. It is the relatively neglected, undernourished, and underachieving subset of research on learning which I call research on teaching that I shall deal with here.

The Stanford Center for Research and Development in Teaching is devoted to this problem area. In its conceptual framework, teaching, or the behaviors and characteristics of teachers, stands at the center. This domain contains variables that serve as both independent and dependent variables in the Center's program. When the teachers' behaviors and characteristics serve as independent variables, and the pupils' learning and behaviors serve as the dependent variables. then we have research on teacher effectiveness, or, more neutrally, research on teacher effects. On the other hand, when teachers' behaviors and characteristics serve as the dependent variables, and teacher education programs and procedures serve as the independent variables, then we have research on teacher education. Taking both research on teacher education and research on teacher effects as our domain, we have research on teaching, as it is understood in the Stanford Center. My subject today is the Center's program of research and development on instructional methods, and my procedure will be to work from the past to the present, from the general to the specific, and then to try for a look at where we are going. When I get down to specifics, I shall be talking primarily about research in which I am involved, rather than all of the research underway in our Center.

But first, let us look at where research on teaching has been. As the behavioral sciences go, it has a respectably long history but a regrettably inglorious one. Research on teaching has been going on almost as long as research on learning. Some studies were made in the 1910's and 1920's, and quite a few were made during the 1930's. By the early 1950's, substantial

reviews and bibliographies of research on teaching began to appear. And during the last decade, the flow of research on teaching has indeed become significant. But the early years did not pay off in solid, replicable, meaningful results that had considerable theoretical or practical value. Positive and significant results were seldom forthcoming, and they survived replication even less often. The research yielded many findings that did not make sense, that did not hang together in any meaningful way.

Under these conditions, as Kuhn (1962) has pointed out, research workers are impelled to re-examine their first principles, the paradigms by which they guide their efforts. The model problems and solutions of the community of researchers on teaching were accordingly subjected to more and more reappraisal. Licking the wounds inflicted by their negative results, researchers on teaching built up a modest literature of new conceptual frameworks, approaches, and paradigms for research on teaching.

To illustrate, let me refer to one of the dominant paradigms that even today leads many discussions and research projects into the wilderness. This is the paradigm that says that what we need above all, before we can select and train better teachers, is the criterion of teacher effectiveness. Here is one example of that kind of approach:

> The lack of an adequate, concrete, objective, universal criterion for teaching ability is thus the primary source of trouble for all who would measure teaching. One typical method of attack used in rating scales is to compile a list of broad general traits supposedly desirable for teachers, with respect to which the rater passes judgment on each teacher. This amounts to an arbitrary definition of good teaching, which is subjective and usually vague, but it does not necessarily lead to an identification of it. Only if the traits themselves can be reliably identified can their possessor be identified as a "good teacher" according to the definition laid down in the scale. Even when the scale is made quite specific, relating not to general traits but to concrete procedure, the fundamental difficulty remains, that there is no external and generally accepted criterion against which the scale can be validated to establish the significance of its items (Walker, 1935, pp. x-xi).

This kind of writing implies that there is some magic variable that applies to all of teaching, for all pupils, at all grade levels, in all subject matters, and in all objectives. The phrase "the criterion of teacher effectiveness" betokens a degree of generality that has seldom been found in any branch of the behavioral sciences. It also reflects the mistaken notion that such a criterion, largely a matter of values, can be established on the basis of scientific method alone.

The so-called criterion problem misled a whole generation of researchers on teaching, embroiled them in endless and fruitless controversy, and lured them into hopelessly ambitious attempts to predict teacher effectiveness over vast arrays and spans of outcomes, teacher behaviors, time intervals, and pupil characteristics, all on the basis of predictive variables that had only the most tenuous theoretical justification in the first place. It is little wonder that, when Berelson and Steiner (1964) dealt with the subject of teachers' behaviors and characteristics in their inventory of scientific findings in the behavioral sciences, they dismissed the "large number of studies" with the single dismal sentence that "there are no clear conclusions" (p. 441).

If the global criterion approach had proved to be sterile, what was the alternative? The answer was to take the same path that more mature sciences had already followed: If variables at one level of phenomena do not exhibit lawfulness, break them down. Chemistry, physics, and biology had, in a sense, made progress through making finer and finer analyses of the phenomena and events they dealt with. Perhaps research on teaching would reach firm ground if it followed the same route.

Apparently, a number of students of the problem had this general idea at about the same time. In 1962, writing my chapter on paradigms for the Handbook of Research on Teaching, I coined the term, "micro-criteria" of effectiveness. As I said in that chapter:

> ...One solution within the "criterion-of-effectiveness" approach may be the development of the notion of "micro-effectiveness." Rather than seek criteria for the over-all effectiveness of teachers in the many, varied facets of their roles, we may have better success with criteria of effectiveness in small, specifically defined aspects of the role.... a sufficient number of laws applying to relatively pure aspects of the

> teacher's role, if such laws could be developed, might eventually be combined, ... to account for the actual behavior and effectiveness of teachers with pupils under genuine classroom conditions" (Gage, 1963, p. 120).

A group of workers at Stanford University, to which I was to move a few months later, took a similar view. In the Stanford program for training secondary school teachers, Robert Bush, Dwight Allen, and their co-workers adopted what is now known as the technical skills approach. Technical skills are specific instructional techniques and procedures that a teacher may use in the classroom. They represent an analysis of the teaching process into relatively discrete components that can be used in different combinations in the continuous flow of the teacher's performance. The specific set of technical skills adopted in the teacher education program at Stanford may be quite arguable. Indeed, the list of skills has been revised a number of times over the past few years. What is important is the approach -- the attempt to analyze teaching into limited, well-defined, components that can be taught, practiced, evaluated, predicted, controlled, and understood in a way that has proven to be altogether impossible for teaching viewed in the larger chunks that occur over a period of an hour, a day, a week, or a year.

When analyzed-teaching, in the form of technical skills, is made the focus of our concern, we find it possible to do fairly satisfying research both on teacher education and on teacher effects. The satisfaction comes from being able to measure or manipulate relevant independent variables, perform true experiments, or make careful analyses and measure relevant dependent variables.

The idea of technical skills may be illustrated by the terms used in a recent list of such skills. One was called "establishing set," or the establishment of cognitive rapport between pupils and teacher to obtain immediate involvement in the lesson; one technique for inducing a positive set is the use of relevant analogies. A second technical skill is that called "establishing appropriate frames of reference," or points of view. A third technical skill is that of "achieving closure," or pulling together major points, linking old and new knowledge, at appropriate points within a teaching episode as well as at the end. A fourth technical skill is that of "using questions" in such a way as to elicit the kinds of thought-processes and behaviors desired, such as simple recall, or concept formation, or evaluation. Other technical skills

are those in "recognizing and obtaining attending behavior," "control of participation," "providing feedback," "employing rewards and punishments," and "setting a model."

These technical skills into which important aspects of the teaching job have been analyzed are not merely the subjects of lectures and discussions in the teacher education program. Rather, they form the basis for the intern's practice teaching prior to his entrance into actual classrooms. This procedure, well known by now as "micro-teaching," consists in getting the trainee to teach a scaled-down teaching exercise. It is scaled down in terms of time because it lasts only five to ten minutes. It is scaled down in terms of class size, because he teaches a group of not more than five pupils, who are brought in and paid to serve as pupils in the micro-teaching clinic. It is scaled down in terms of the task, since the trainee attempts to perform only one of the technical skills in any single micro-teaching session. The sessions are recorded on video tape, and the trainee gets to see and hear himself immediately after the session. While he looks at and listens to himself, he receives criticisms and suggestions from supervisors trained to be both perceptive and tactful. Then he "re-teaches" the same lesson to a new small group of pupils in an attempt to improve on his first performance of the specific technical skill that is his concern in that session.

Obviously, the general idea is subject to many variations. The size of the class can be manipulated, the number of trainees teaching a given group of children can be increased, the duration of the lessons can be lengthened, and the nature of the teaching task can be made more complex so as to embrace a group of technical skills in their real-life combinations. But the idea of analyzing teaching into technical skills remains the heart of the method and provides its power as a paradigm for research.

The research on micro-teaching and technical skills in the Stanford teacher education program has taken the form of experiments in which various procedures for feedback to the trainee are manipulated. Professors Dwight Allen and Frederick McDonald have organized a program of research on variables hypothesized to influence the learning of the technical skills of teaching. Their independent variables fall into three categories: practice variables, feedback variables, and demonstration variables. A practice variable may consist in micro-teaching versus teaching in an actual classroom. A feedback

variable may be the positive or negative character of the feedback, or the mediation of the feedback by another person rather than the trainee himself. Finally, a demonstration variable may take the form of symbolic demonstration, consisting of written or spoken words, or perceptual demonstration, consisting in either live or video taped portrayals of the desired behavior; and each of these can consist either of self-modeling or modeling by others. Other independent variables have been identified, such as the timing of a reinforcement, the amount of practice, and the amount of feedback.

This condensed description of the Allen-McDonald research program can suffice to illustrate the use of the analytic approach to research on teacher education. Their research takes the form of true experiments in which subjects are randomly assigned to different values of the independent variable.

I should like to turn now to an example of the way in which the technical skills approach can be applied to the study of teacher effects. This research has dealt with a technical skill that I call "explaining," or the skill of engendering comprehension -- usually orally, verbally, and extemporaneously -- of some process, concept, or generalization. Explaining occurs in all grade levels and subject matters, whether it is a fifth-grade teacher explaining why the time in New York differs from that in San Francisco or a geologist explaining how the ice age may have been caused by volcanic eruptions. Everyday observation tells us that some people explain aptly, getting to the heart of the matter with just the right terminology, examples, and organization of ideas. Other explainers, on the contrary, get us and themselves all mixed up, use terms beyond our level of comprehension, draw inept analogies, and even employ concepts and principles that cannot be understood without an understanding of the very thing that is being explained. To some of us, it has seemed that explaining comes very close to being the inner essence of instruction, so that when a teacher is attempting to explain proportionality to his geometry class or irony to his English class, he is behaving more purely as a teacher than when he is attempting, say, to motivate, promote discussion, or maintain discipline. At any rate, we have made some studies of explaining ability in the attempt to determine some of the characteristics of effective explanations.

The first study was made in the micro-teaching clinic at Stanford during the summer of 1965 by J. C. Fortune, R. F. Shutes, and me. We attempted to determine the generality of explaining ability, that is, the degree to which the ability to explain one topic was correlated with the ability to explain another topic, and the degree to which the ability to explain a topic to one group of pupils on one occasion was correlated with the ability to explain the same topic to another group of pupils on another occasion. We also were able to design the study so as to determine the degree to which there was generality over both pupils and topics, or the degree to which the ability to explain one topic to one group of pupils on one day correlated with the ability to explain another topic to another group of pupils on another day. Because there were only 60 pupils to be shared in groups of 5 among approximately 40 interns in the micro-teaching clinic, the design became quite complex in order to avoid having any intern teach the same topic to the same group of pupils more than once and to avoid having the same group of pupils receive an explanation of the same topic more than once. Accordingly, the 40 social studies interns -- and we chose to work with the social studies interns only because there were more of them than any other kind of intern -- were divided into five clusters of eight interns each. The lectures dealt with 20 different topics, each consisting of an Atlantic Report from the Atlantic Monthly. The correlations that we obtained were thus medians of five correlations, each based on four, six, six, six, and eight interns, respectively.

The index of lecture effectiveness, or what I would like to call the micro-criterion of teacher explaining ability, was the pupils' mean score on a ten-item test of their comprehension of the main ideas of the lecture, which was presented by each intern in 15 minutes under somewhat standardized conditions. This mean score was adjusted for the mean ability of the pupils in the given group as measured by their scores on all of the other topics. Similarly, any given mean score was adjusted for the difficulty of the topic as measured by the mean score of all groups of pupils on that topic. Thus, the variance of the adjusted mean posttest comprehension scores was attributable not to the ability of the pupils or the difficulty of the topic, but rather to the differences among the teachers. We then investigated the question of the various kinds of generality by determining the median intercorrelations among

the various means. The upshot of this part of the study was that generality over topics was non-existent, and generality over groups was about .4. In other words, the interns were moderately consistent in their ability to explain the same topic to different groups on different occasions, but they were not consistent in their ability to explain different topics.

This study also dealt with the correlations between explaining effectiveness and the pupils' ratings of various aspects of the explanations. The pupils rated the interns' performance with respect to twelve items, such as clarity of aims, organization of the lesson, selection of material, and clarity of presentation. To us it seemed that some of these dimensions should correlate more highly with explaining ability than others. In particular, we hoped that such discriminant validity would be manifest in the form of a higher correlation between the mean rating of the lecture for "clarity of presentation" than for any of the other items of the Stanford Teacher Competence Appraisal Guide. Our hope was borne out; the correlation of the adjusted mean posttest comprehension scores with pupils' ratings of "clarity of presentation" was .56, higher than that with any of the other rating scale items. This result seems to us to support the validity of both the index of lecture effectiveness and the mean ratings by the pupils.

During the school year 1965-1966, I was able, in collaboration with Barak Rosenshine, to undertake a replication and extension of this study in the public schools. Because there was no lack of students in the high school classes, taught by their own teachers, we did not become involved in the complexities of design necessary in the micro-teaching clinic. To put it very briefly, we got 40 eleventh-grade social studies teachers each to deliver a 15-minute lecture on an Atlantic Report on Yugoslavia taken from the Atlantic Monthly. The teachers had been given the article several days in advance, and had been told to prepare a lecture that would enable their pupils to answer a ten-item multiple-choice test of comprehension of the article's main ideas. To guide them in preparing their lecture, they were given five of the multiple-choice questions that would be asked, while the other five questions were withheld. After the 15-minute lecture, in which the teachers were permitted to use the blackboard but no other aids, their students took the ten-item test. They also rated the teacher's lecture on items similar to those I have already described. The next day, the same teachers and classes did the same things, except that the subject matter was an Atlantic Report on Thailand; again the

teachers had been given five of the ten items as a guide to the kind of lecture that they should prepare and had been told to focus on the explanation of the major ideas, concepts, and principles brought out in the article, which constituted the curriculum for this bit of teaching. On the third day, the classes heard a third lecture, one that was the same for all classes, a tape recorded 15-minute lecture on Israel, a verbatim reading of an Atlantic Report, and then the pupils again took a ten-item test based on that article.

The class mean on the Israel test was used to adjust the class means on Yugoslavia and Thailand for between-class differences in ability. Our reasoning was that the score on such a test of comprehension of a uniform lecture would be more useful in controlling relevant kinds of ability than would the usual scholastic aptitude test. The class means on Yugoslavia and Thailand were also adjusted for teacher differences in the content-relevance of the lecture, as determined by scoring the transcript of the lecture for relevance to the ten items on the comprehension test. We then assumed that the variance that still remained in the adjusted comprehension test means of the classes would reflect differences between the teachers in what we were concerned with, namely, the intellectual style and process of the teacher's lecture. In this study, the teacher's adjusted effectiveness index on Yugoslavia correlated .47 with his effectiveness on Thailand; i.e., there was considerable generality of effectiveness over topics, even after student ability and content relevance had been partialed out.

It should be noted that we were using the micro-teaching idea in this investigation. The teaching was restricted to just one aspect of the teacher's role, namely, ability to explain the current social, political, and economic situation in another country. The curriculum was also scaled down. We also used another major feature of the micro-teaching clinic, the videotape recorders which made it possible for us to study the teacher's behavior, verbal and nonverbal, at leisure.

One major question was that of whether our criterion, or micro-criterion, of teacher effectiveness in explaining, namely, the mean comprehension score of the pupils, adjusted for both mean pupil ability and content-relevance, contained variance that would be manifested in something about the lecture that was visible or audible. In other words, was there some difference between good and poor explanations that was worth trying to analyze? So we picked two

lectures on Yugoslavia that were extremely high on our index of effectiveness and two that were extremely low. We had a group of eight judges read the article on Yugoslavia and take the comprehension test, and then watch and listen to all four of these lectures. Then the judges ranked the lectures in terms of perceived effectiveness in engendering comprehension as measured by the ten-item test. It turned out that the judges' postdictions were quite significantly more accurate than could have been expected on the basis of chance, and we were accordingly reassured that our micro-criterion was indeed reflected in something that could be seen or heard in the lecture.

But the major concern of this investigation was to determine the cognitive and stylistic correlates of the lecture's effectiveness. For this purpose, we used extreme groups to minimize the labor of scoring a host of variables about which we had no great conviction. So the ten most effective explanations on Yugoslavia were identified and also the ten least effective. From these, we chose at random five of the most effective and five of the least effective. Then, groups of judges and content analysts worked over the transcripts of the lectures, scoring and rating them on a host of variables. Some of these were sentence fragments, the average sentence length, the number of prepositional phrases per sentence, and so on. Other variables dealt with the number of self-references by the teacher, or with various aspects of syntax, or instructional set, familiarization, uses of previous knowledge, mobilizing sets, attention focusing procedures, organization, emphasis, amount of repetition and redundancy, the number of words per minute, and so on.

The variables that discriminated between the five best and the five worst lectures on Yugoslavia were then tried out on the other set of five best and five worst lectures on Yugoslavia to see if they still discriminated. Those that survived this first cross-validation were then tried out on the best and worst lectures on Thailand. At the last accounting, two characteristics of the lectures had survived this kind of validation and cross-validation procedure. These variables were what we are calling "explaining," or the degree to which the teacher describes the how, why, or effect of something, and the "rule-eg-rule" pattern, or the degree to which the teacher states a generalization, gives examples of it, and then summarizes a series of illustrations at a higher level of generality than the illustrations themselves. These two variables not only

seem to be valid in our data but also are reliably rated by independent judges. Nonetheless, these must not be considered to be firmly established findings; they are merely examples of the kinds of conclusions to which research of this kind can lead.

Currently, we are in the process of scoring all of the explanations on all of the variables that appear to hold any promise, and we will then undertake studies of the complete correlation matrices involving not only the indices of explanation effectiveness, but also all of the characteristics of the explanations, and the ratings of the lectures by the pupils who heard them. Such a complete correlational study will throw light on the consistency from one lecture to another of the indices of lecture effectiveness, and the stylistic characteristics of the lectures, and also their intercorrelations.

What I have been describing is of course a correlational study. Along with its advantages in permitting the exploration of a wide variety of possible correlates of explaining ability as they occur under fairly normal conditions, it also has the disadvantage of making causal interpretations hazardous. For this reason, studies of this kind ought to proceed fairly rapidly into experiments in which the different ways of explaining will be based, at least in part, on leads obtained from our correlational studies.

Such experimental research may lead toward quite novel methods of teaching that could never be developed on the basis of studies of teaching the way teaching is. Stolurow has contrasted the approach of "modeling the master teacher" with that of "mastering the teaching model." The first approach is that of studying the most effective teachers we can find in order to find out how they behave and what they are like so that we can attempt to produce more teachers like them. Many research workers see little promise in this approach. They recommend that we undertake instead to develop wholly new models of the teaching process designed for optimal effectiveness regardless of their similarity to the way teaching now goes on in the normal classroom.

The teaching model that many advocates of this approach have in mind is that of programmed instruction, particularly computer-assisted programmed instruction. As Suppes and Atkinson and others have described this revolutionary undertaking in research and development on instruction, it holds out great promise indeed. Before too long, the annoying problems in the hardware will have been solved. After a somewhat longer time, we may expect substantial

and well-validated programmed curricular materials to have been developed in all the subject matters and grade levels. As one who has seen the highly developed installations at the Brentwood School in East Palo Alto, California, I must share the optimism of Suppes and Atkinson, and other developers of computer-assisted instruction.

Their very success, or coming success, raises problems for the kind of instructional methods with which we have been concerned today. On superficial examination, at least, certain major problems of ordinary classroom teaching seem to be clearly surmounted by computer-assisted instruction. For example, the problem of the cognitive complexity in the teacher's task, of how the teacher can say just the right thing at the right time to develop a concept or formulate a theory, is apparently well handled, at least in principle, by computer-assisted instruction. Its programs can be worked out and tried out in meticulous detail, well in advance, at leisure, by the most skilled curriculum experts in the land, and then made available in all their subtlety and complexity to every teacher who uses the program. Another major problem in the ordinary classroom is that of individualizing instruction. No matter how we group our pupils between schools, within schools, or within classrooms, we still have the problem of adjusting the rate and direction of the teaching and learning process to the needs and abilities of the individual pupil; here again, at least in principle, computer-assisted instruction seems at first glance to have the better of the live teacher.

While pondering these problems, I got some help from a restatement of the idea of individualized instruction in a recent paper by Philip Jackson. As he put it,

> Individualizing instruction, in the educator's sense, means injecting humor into a lesson when a student seems to need it, and quickly becoming serious when he is ready to settle down to work; it means thinking of examples that are uniquely relevant to the student's previous experience and offering them at just the right time; it means feeling concerned over whether or not a student is progressing, and communicating that concern in a way that will be helpful; it means offering appropriate praise, not just because positive reinforcers strengthen response tendencies, but because the student's performance is deserving of human admiration; it means, in short, responding <u>as</u> an individual <u>to</u> an individual (Jackson, 1966).

Individualization in this sense is much more than allowing for differences in speed of moving through a program or providing different branches or routes through the material.

Jackson's analysis of this kind of limitation in computer-assisted instruction should be placed alongside of the indications by Suppes (1966) that tutoring and dialog,which are higher levels of instruction than drill-practice, are still well in the future, as capabilities of computer-based instruction. Hence, any fears about the rapid obsolescence of live teachers, even where narrowly defined cognitive objectives are concerned, are quite unwarranted. That is, there will still be a need for teachers to use the kinds of technical skills, including explaining, with which the analytic approach being developed at Stanford and elsewhere is concerned. We shall have to continue to grapple with the problems of cognitive complexity and individualization through the medium of the live, human, teacher even in the realm of the well-formulated cognitive objectives. And there will always be the indispensable role of teachers in assisting pupils in attaining various kinds of affective and social learnings in the classroom.

Accordingly, group discussions, role playing, teaching for divergent thinking, as well as the technical skills I have already mentioned, are all the subject of various research and development projects now under way in the Stanford Center for Research and Development in Teaching. We are also looking at the way in which the organizational context influences the teacher's choices among ways to teach. And, in one of our projects, entitled "The Teacher in 1980," we are looking at the way in which new curriculum developments, television and other technical aids, computer technology, and new organizational schemes in the schools will affect the teacher's role in the foreseeable future.

In conclusion, let me refer again to what I see as one basic new theme in the research and development in teaching that is now under way at Stanford and elsewhere. If it were necessary to sum it up in one word, my word would be <u>analysis</u>, breaking down the complexities that have proven to be so unmanageable when dealt with as a whole. We are no longer crippled by the notion that because there is one word called teaching, there is one, single, over-all criterion of effectiveness in teaching that will take essentially the same form wherever teaching occurs. Even if none of the present analyses

that we have now proves to be viable, they will not be replaced by the old, global, conceptually impossible, complex variables that I see as the reason for the fruitlessness of so much of research on teaching in the past. Instead, they will be replaced by other analyses of teaching, perhaps even finer analyses, until we get the sets of lawful relationships between variables that will mark the emergence of a scientific basis for the practice of teaching. It may well be that a 15-minute explanation of a 5-page magazine article is still too large a unit of teaching behavior to yield valid, lawful knowledge. It may well be that the mean score on a 10-item test of comprehension, adjusted for student ability and content relevance of the lecture, is still too large and complex a dependent variable. But, compared with the massive, tangled, and unanalyzable units that have typically been studied in the past -- in research on the lecture method, the discussion method, and class size, for example -- such units seem precise and manageable indeed. And eventually, of course, we shall have to put teaching back together again into syntheses that are better than the teaching that goes on now. I think it would be safe to say that there is now some hope of our being able to develop a scientifically grounded set of answers to every teacher's central question, "What should I do in the classroom?"

References

Berelson, B, & Steiner, G. Human behavior: An inventory of scientific findings. New York: Harcourt, Brace, & World, 1964.

Gage, N. L. Paradigms for research on teaching. In N. L. Gage (ed.), Handbook of research on teaching. Chicago: Rand, McNally, 1963. Pp. 94-141.

Jackson, P. W. The teacher and the machine: Observations on the impact of educational technology. (Mimeographed. Chicago: The author, University of Chicago, 1966).

Kuhn, T. S. The structure of scientific revolutions. Chicago: University of Chicago Press, 1962.

Suppes, P. The uses of computers in education. Scientific American, 1966, 215(3), 207-220.

Walker, Helen M. Preface. In Lancelot, W. H., et al., The measurement of teaching efficiency. New York: Macmillan, 1935. Pp. ix-xiv.

JUL 1 8 1967